Interfaces for Wide Area Information Servers

Brewster Kahle

Thinking Machines Corporation

Talk Overview

- New problems for computer interfaces
- WAIS servers:
  - DowQuest
  - SWAIS
  - GWAIS
  - XWAIS
  - WAIStation
  - Rosebud
- Conclusion

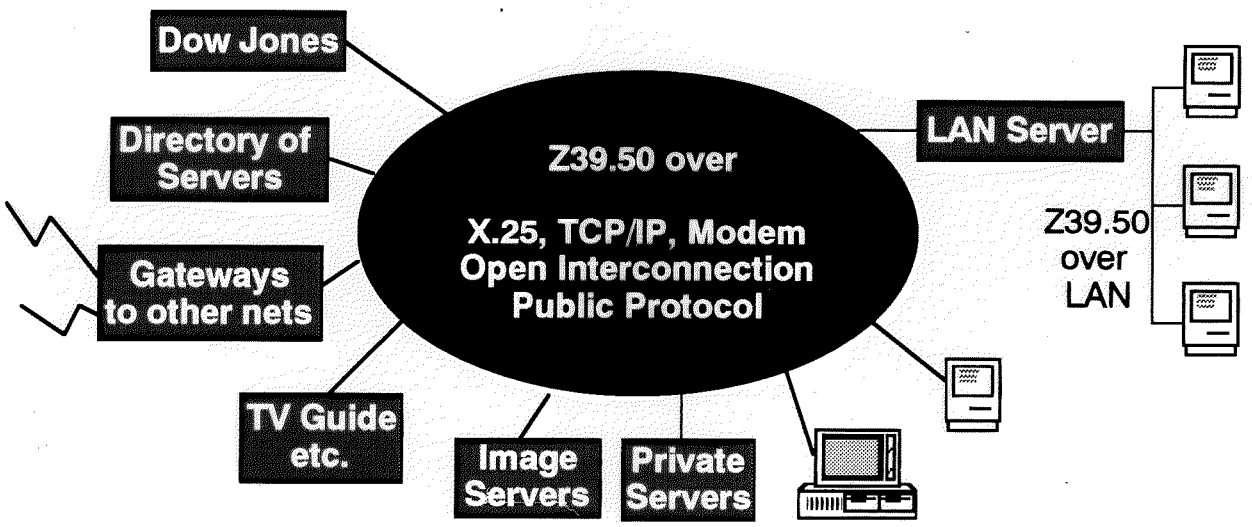
New Problems

- Your machine will be spending your money.
- The information on your computer will be constantly changing.
- Browsing online information will become common.
- Your machine will know more about you than your spouse.

Technology Limits

- Screen size and resolution
- Communications speed
- User communities' acceptance

WAIS Architecture



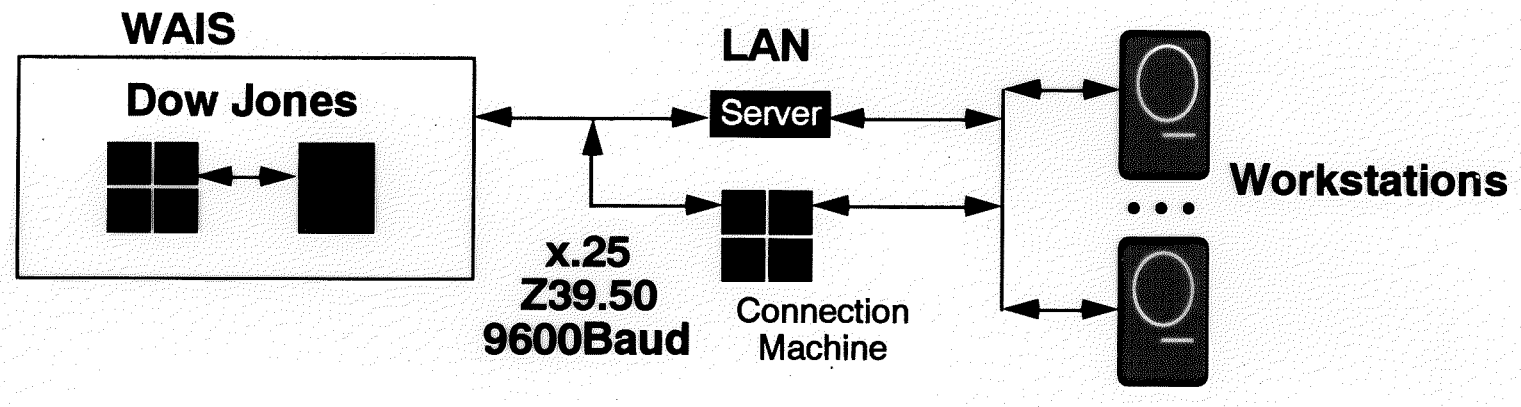
Users Needs:

- Selecting Servers
- Answering Questions
- Organizing Responses

Architecture Issues:

- Scalability
- Security
- Business model for servers
- Reliable Access

Office Demonstration Structure



Operations:
- Archiving
- Queries
- Retrieval

IR Type:
- Broadcast
- Query by Example

Databases:
- Wall St Journal
- Barron's
- 400 Business Mags

CM:

Operations:
- Queries

IR Type:
- enhanced relevance feedback

Databases:
- DowVision and memo's, mail, word processor files

Mac:

Operations:
- Human Int
- Retrieval
- Queries
- "Caching" Docs
- User Profiles

IR Type:
- Query by example

Databases:
- Personal Text
- Cached data

DowQuest Startup and Question

```
quake.think.com
      ALL RIGHTS RESERVED

CONGRESS NEARS FINAL APPROVAL
OF JOBLESS-BENEFITS EXTENSION
BUT SENATE VOTE ISN'T
VETO-PROOF, SEE //NEWS.
ENTER QUERY
                                DOWQUEST
              Copyright (C) 1991 Dow Jones & Company, Inc.

                                * * * * *
DowQuest searches articles from more than 350 sources of current business
information, including The Wall Street Journal, Barron's, Business Week,
Fortune and Forbes.  All articles are copyrighted by their publisher.

                                * * * * *
TO VIEW INSTRUCTIONS or copyright notices, press (Return).
For additional information, type HELP at any point while using DowQuest.

TO BEGIN A SEARCH: Enter several words that are relevant to the subject
you are interested in and press (Return).

ENTER QUERY:
   Thinking Machines and IBM
```

DowQuest Search Response

quake.think.com

```
ENTER QUERY:
DOWQUEST                          STARTER LIST              HEADLINE PAGE 1 OF 4

  1 Technology:  IBM, Thinking  Machines Plan  Joint Projects
    WALL STREET JOURNAL , 09/24/91  (657 words)

  2 Thinking Machines Wins 'Edge'
    BOSTON GLOBE, 03/25/91  (960 words)

  3 Fast Thinking Fuels Computer Rivalries
    BOSTON BUSINESS JOURNAL, 07/01/91  (897 words)

  4 Technology:  Thinking Machines Retakes  Title for . . .
    WALL STREET JOURNAL , 06/05/91  (246 words)

                                  * * * * *
TO IMPROVE A SEARCH: Type SEARCH and the numbers of up to three articles
                       that are examples of your subject.
TO VIEW AN ARTICLE:  Enter article number and press (Return).
                     Enter BEST and article number to see "best" section.
TO PRINT ARTICLES:   Type PRINT and the article numbers.

PRESS (RETURN) FOR MORE HEADLINES. TYPE HELP FOR MORE INFORMATION.
```

DowQuest Document Screen

quake.think.com

```
PRESS (RETURN) FOR MORE HEADLINES. TYPE HELP FOR MORE INFORMATION.
DOWQUEST                        ARTICLE 1                    PAGE 1 OF 5

    STORY ID: 0000167878WJ
  1 Technology:
  2  IBM, Thinking
     Machines Plan
     Joint Projects


    WALL STREET JOURNAL (J) 09/24/91
    Copyright (c) 1991 Dow Jones & Company, Inc.

  3 By Paul B. Carroll
     Staff Reporter of The Wall Street Journal

  4   International Business Machines Corp. reached a development
    agreement with Thinking Machines Corp. that should strengthen IBM's
                                  * * * * *
TO IMPROVE A SEARCH: Look for paragraphs that are examples of the subject
  you are interested in. Type SEARCH and up to three paragraph numbers.
TO VIEW MORE OF THIS ARTICLE: Press (Return).
TO RETURN TO THE HEADLINES:   Type /M (Return).
```

DowQuest Help Screen

quake.think.com

```
TO RETURN TO THE HEADLINES:   Type /M (Return).
When you entered an article number, DowQuest displayed the article with
each paragraph numbered.You can improve your quest by indicating which
paragraphs are good examples of what you are looking for. Type SEARCH
and up to three paragraph numbers. DowQuest will retrieve another
16 articles based on these example paragraphs.

If the entire article is a good example, type SEARCH without numbers.

You can also:

-- Review more of the article. Press (Return) to go forward one page.
   Type /R to go back one page.

-- Print the article without the numbers and instructions. Type PRINT.

-- Return to the list of headlines. Enter /M.

-- View another article. Enter the article number.

-- Begin a new quest. Enter /T. To leave News/Retrieval, type DISC.

PRESS (RETURN) TO VIEW ARTICLE AGAIN. THEN ENTER REQUEST.
```

DowQuest: What's New?

- **English Language + Relevance Feedback**

Designer: DowJones
Use: Call DowJones at 609-520-4000

DowQuest: Problems

- Only published information
- TTY interface
- $2/minute (browsing expensive)
- 2400 baud access

SWAIS Source Window

```
                                   SWAIS
SWAIS                          Source Selection                          Sources: 79
  #            Server                          Source                          Cost
 55:   [      cmns.think.com]  patent-sampler                                   Free
 56:   * [microworld.media.mit]  poetry                                         Free
 57:   [ lambada.oit.unc.edu]  rec.cook                                         Free
 58:   [ lambada.oit.unc.edu]  rec.pets                                         Free
 59:   [      cmns.think.com]  risks-digest                                     Free
 60:   [ pit-manager.mit.edu]  rkba                                             Free
 61:   [     quake.think.com]  sample-books                                     Free
 62:   [     quake.think.com]  sample-pictures                                  Free
 63:   [          129.71.11.2]  sorrel-ada-archives                             Free
 64:   [      cmns.think.com]  sun-spots                                        Free
 65:   [  talon.ucs.orst.edu]  supreme-ct                                       Free
 66:   [     quake.think.com]  tmc-library                                      Free
 67:   [  uncvx1.oit.unc.edu]  unc-jobs                                         Free
 68:   [ariel.its.unimelb.ED]  unimelb-research                                 Free
 69:   [     quake.think.com]  unix-manual                                      Free
 70:   [   next2.oit.unc.edu]  unix.FAQ                                         Free
 71:   [  violet.cs.uq.oz.au]  usenet-FAQ                                       Free
 72:   [      cmns.think.com]  usenet-cookbook                                  Free

Keywords: What Yeats poem is about a falcon and falconer?

<space> selects, w for keywords, arrows move, <return> searches, q quits, or ?
```

SWAIS Resulting Documents

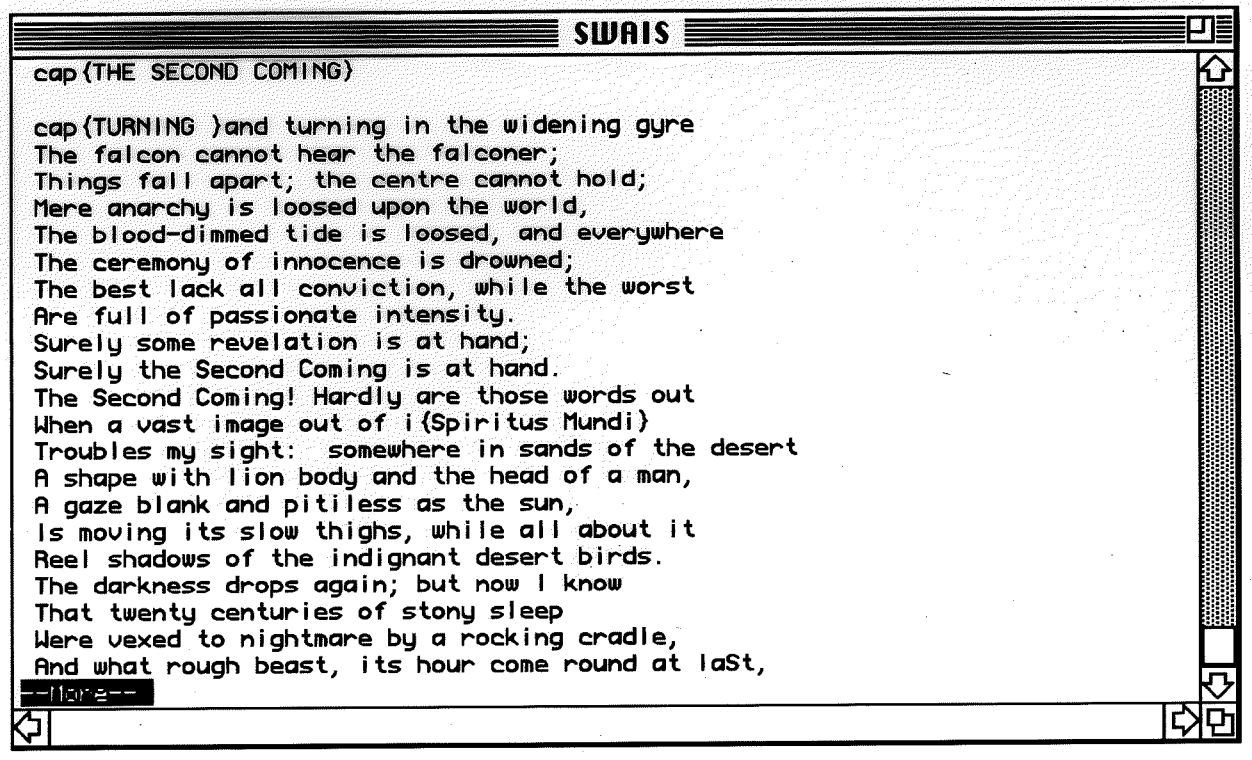
SWAIS

SWAIS Search Results Items: 40

#	Score	Source	Title	Lines
01:	[1000] (	POETRY)	TheSecondComing /tmp_mnt/net/library-dis	24
02:	[1000] (	POETRY)	WanderingsOfOisin /tmp_mnt/net/library-d	1301
03:	[778] (	POETRY)	DramaticPoem /tmp_mnt/net/library-disk/P	807
04:	[778] (	POETRY)	GiftOfHarun /tmp_mnt/net/library-disk/Po	204
05:	[778] (	POETRY)	ShadowyWaters /tmp_mnt/net/library-disk/	48
06:	[778] (	POETRY)	TheFisherman /tmp_mnt/net/library-disk/P	42
07:	[778] (	POETRY)	WarPoem /tmp_mnt/net/library-disk/Poetry	8
08:	[593] (	POETRY)	Epitaph /tmp_mnt/net/library-disk/Poetry	8
09:	[593] (	POETRY)	NineteenNineteen /tmp_mnt/net/library-di	133
10:	[593] (	POETRY)	UnderBenBulben /tmp_mnt/net/library-disk	113
11:	[593] (	POETRY)	UnderSaturn /tmp_mnt/net/library-disk/Po	19
12:	[556] (	POETRY)	ADialogue /tmp_mnt/net/library-disk/Poet	74
13:	[556] (	POETRY)	ADrunkenMan /tmp_mnt/net/library-disk/Po	20
14:	[556] (	POETRY)	AFriendsIllness /tmp_mnt/net/library-dis	9
15:	[556] (	POETRY)	AManYoungAndOld /tmp_mnt/net/library-dis	190
16:	[556] (	POETRY)	AMeditation /tmp_mnt/net/library-disk/Po	7
17:	[556] (	POETRY)	AMemoryOfYouth /tmp_mnt/net/library-disk	31
18:	[556] (	POETRY)	APoetToHisBeloved /tmp_mnt/net/library-d	10

<space> selects, arrows move, w for keywords, s for sources, ? for help

SWAIS Document Window



SWAIS: What's New?

- **Multiple sources**
- **Remote sources**
- **Context sensitive help**
- **Screen based**

Who: John Curran of NSF
Use: Telnet quake.think.com login wais

SWAIS: Problems

- Directory of servers
- Scrolling through many servers

GWAIS Question Window

GWAIS for GNU-EMACS

chocolate cake

usenet-cookbook.s$ | CHOC-CAKE-1(D) USENET Cookbook CHOC-CA

1000	2K	(01/10/90)	CHOC-CAKE-1(D)	USENET Cookbook	CHOC-CA
725	3K	(01/10/90)	CHOC-CAKE-2(D)	USENET Cookbook	CHOC-CA
607	2K	(01/10/90)	CHOKLADKAKA(B)	USENET Cookbook	CHOKLAD
566	2K	(01/10/90)	CHOC-CAKE-4(D)	USENET Cookbook	CHOC-CA
506	2K	(01/10/90)	CHOC-CAKE-5(D)	USENET Cookbook	CHOC-CA
490	3K	(01/10/90)	CHOC-CAKE-3(D)	USENET Cookbook	CHOC-CA
461	3K	(01/10/90)	CHOC-PUDDING-2(D)	USENET Cookbook	CHOC-PUDDI
460	8K	(01/10/90)	CHOC-SLICES(D)	USENET Cookbook	CHOC-SL
447	2K	(01/10/90)	SOURMILK-CAKE(D)	USENET Cookbook	SOURMILK-
431	6K	(01/10/90)	PRINZ-TORTE(D)	USENET Cookbook	PRINZ-T
420	2K	(01/10/90)	CHEESECAKE-4(D)	USENET Cookbook	CHEESECA
420	2K	(01/10/90)	PUMPKIN-CAKE-2(D)	USENET Cookbook	PUMPKIN-CA
416	2K	(01/10/90)	CHOC-PUDDING-1(D)	USENET Cookbook	CHOC-PUDDI
413	2K	(01/10/90)	BLKFOREST-PIE(D)	USENET Cookbook	BLKFOREST
412	2K	(01/10/90)	FUDGE-1(D)	USENET Cookbook	FUD
412	1K	(01/10/90)	BUTTERNUTS(D)	USENET Cookbook	BUTTER
403	3K	(01/10/90)	CHEESECAKE-6(D)	USENET Cookbook	CHEESECA
401	2K	(01/10/90)	POUND-CAKE-2(C)	USENET Cookbook	POUND-CA
399	1K	(01/10/90)	TRUFFLES-3(D)	USENET Cookbook	TRUFFL
397	4K	(01/10/90)	ZUCCOTTO(D)	USENET Cookbook	ZUCC
395	1K	(01/10/90)	CHEESECAKE-3(D)	USENET Cookbook	CHEESECA
392	3K	(01/10/90)	CHEESECAKE-8(D)	USENET Cookbook	CHEESECA
387	4K	(01/10/90)	CHEESECAKE-1(D)	USENET Cookbook	CHEESECA
382	3K	(01/10/90)	CHOC-NUT-TORTE(D)	USENET Cookbook	CHOC-NUT-T
382	2K	(01/10/90)	CHOC-CHIP-3(D)	USENET Cookbook	CHOC-CH
380	3K	(01/10/90)	CARROTCAKE-1(D)	USENET Cookbook	CARROTCA
379	2K	(01/10/90)	SABBATH-STEW(M)	USENET Cookbook	SABBATH-
375	2K	(01/10/90)	CHOC-CHIP-2(D)	USENET Cookbook	CHOC-CH
374	2K	(01/10/90)	CHOC-PIE-1(D)	USENET Cookbook	CHOC-P
374	2K	(01/10/90)	CARROTCAKE-2(D)	USENET Cookbook	CARROTCA

Found 40 documents.

GWAIS: What's New?

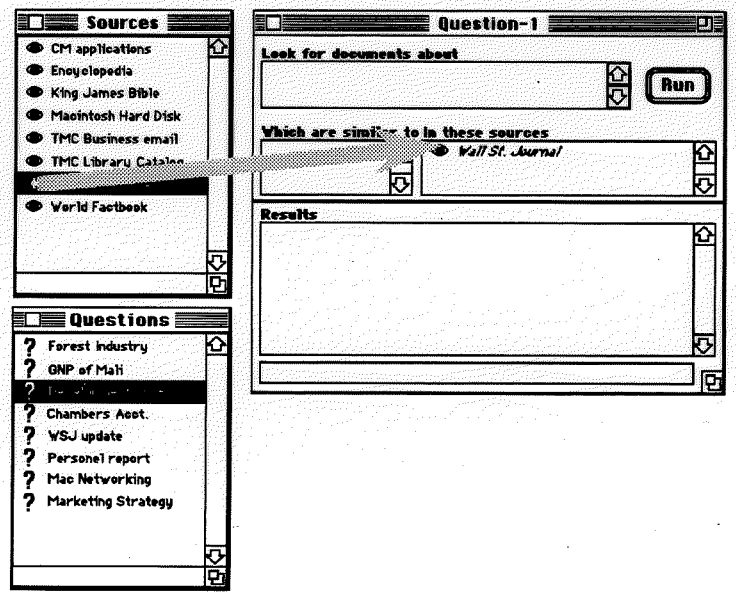
- **Relevance feedback of documents or paragraphs**
- **Integrated into an editor (gnu-emacs)**
- **Screen-based, but multi-window**
- **Non-modal**
- **Supports creation of private sources via indexer**

Designer: Jonathan Goldman of Thinking Machines
Use: anonymous ftp /public/wais/wais-8-*@think.com

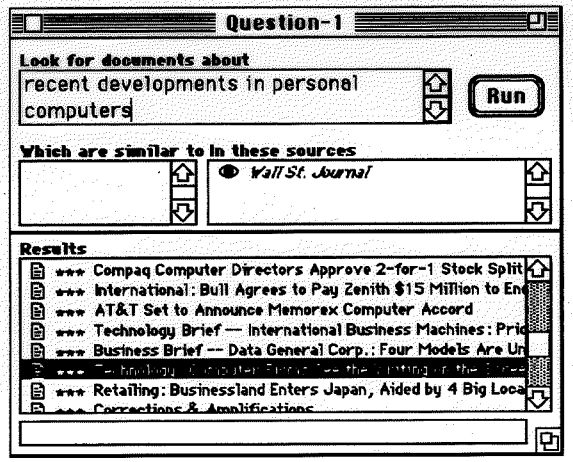
GWAIS: Problems

- Limited to emacs users
- Directory of Servers
- Moving between windows

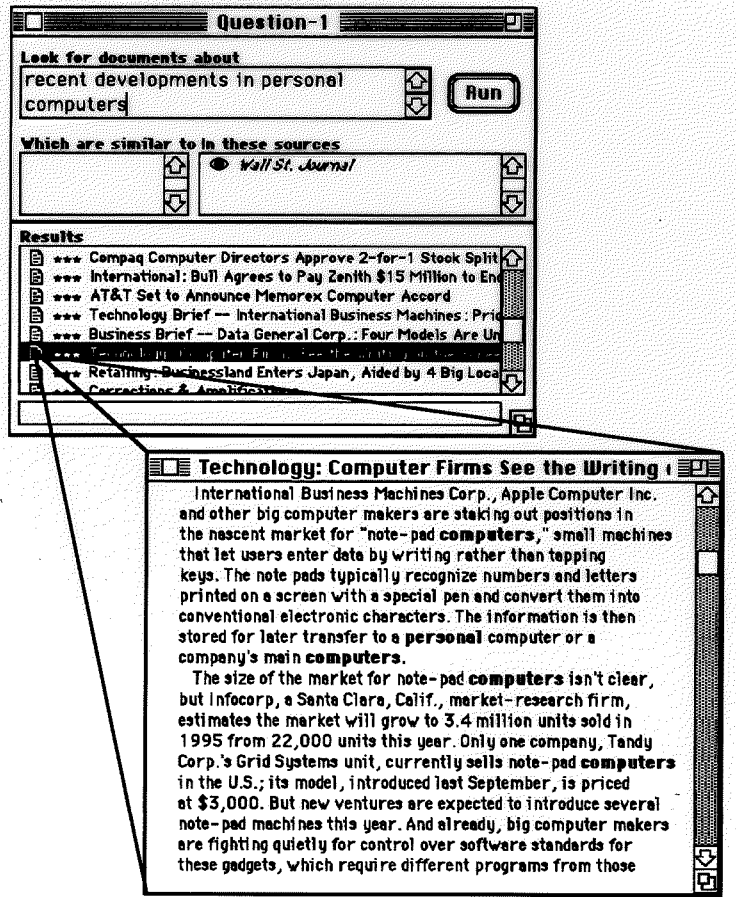
WAIStation #1: Select Source



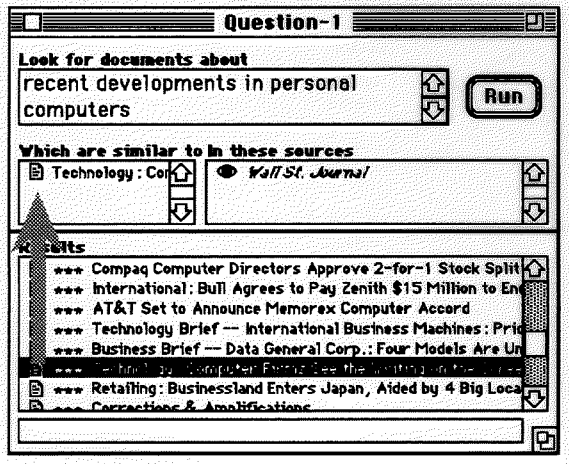
WAIStation #2: Ask Question



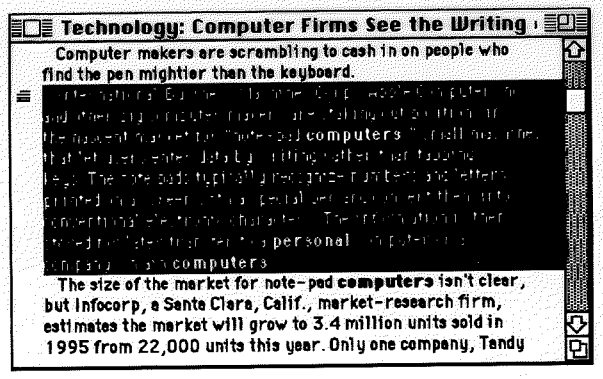
WAIStation #3: Retrieve Documents

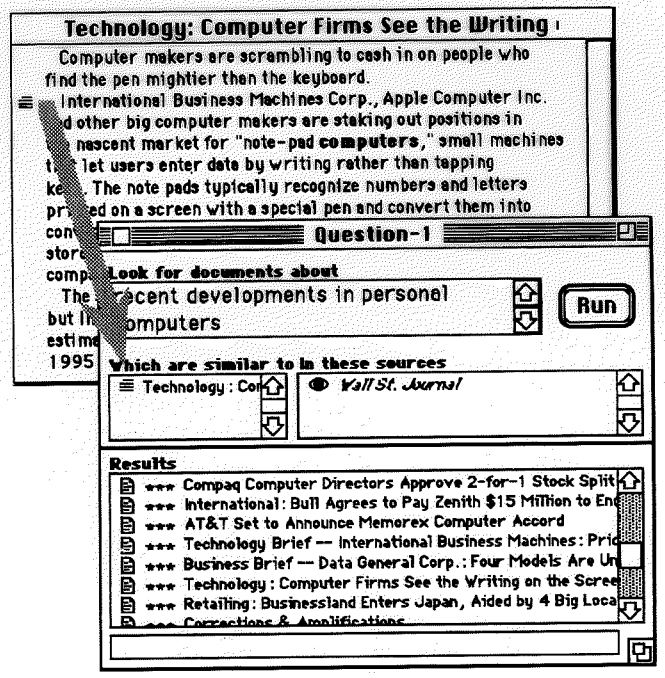


WAIStation #4: Relevant document

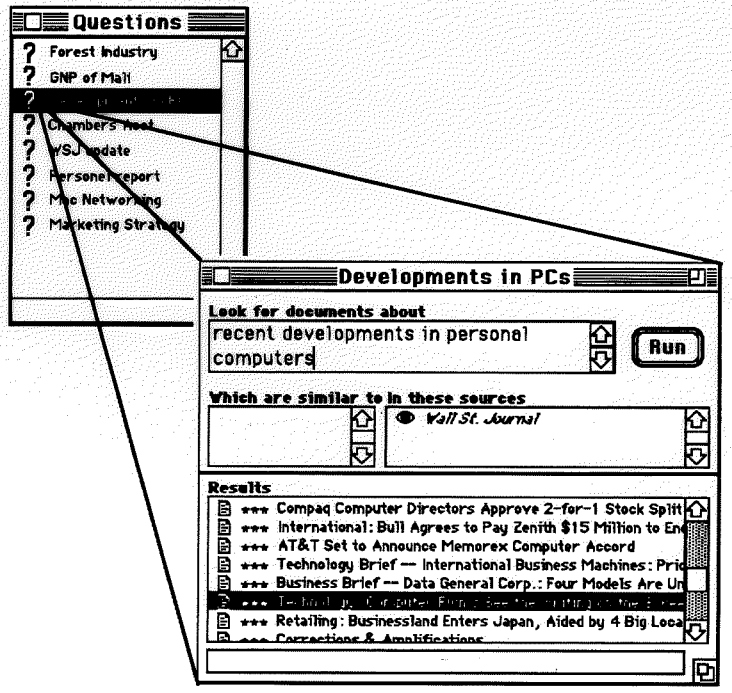


WAIStation #5: Relevant Paragraph

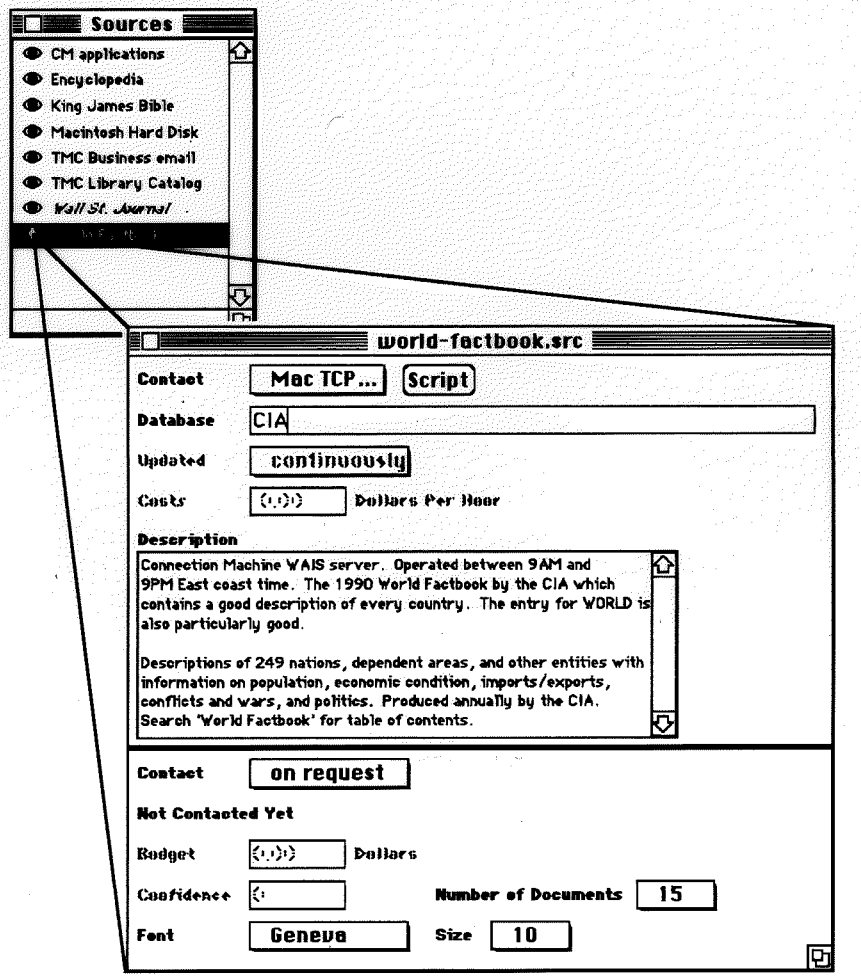




WAIStation Saved Question



WAIStation Source Window



WAIStation: What's New?

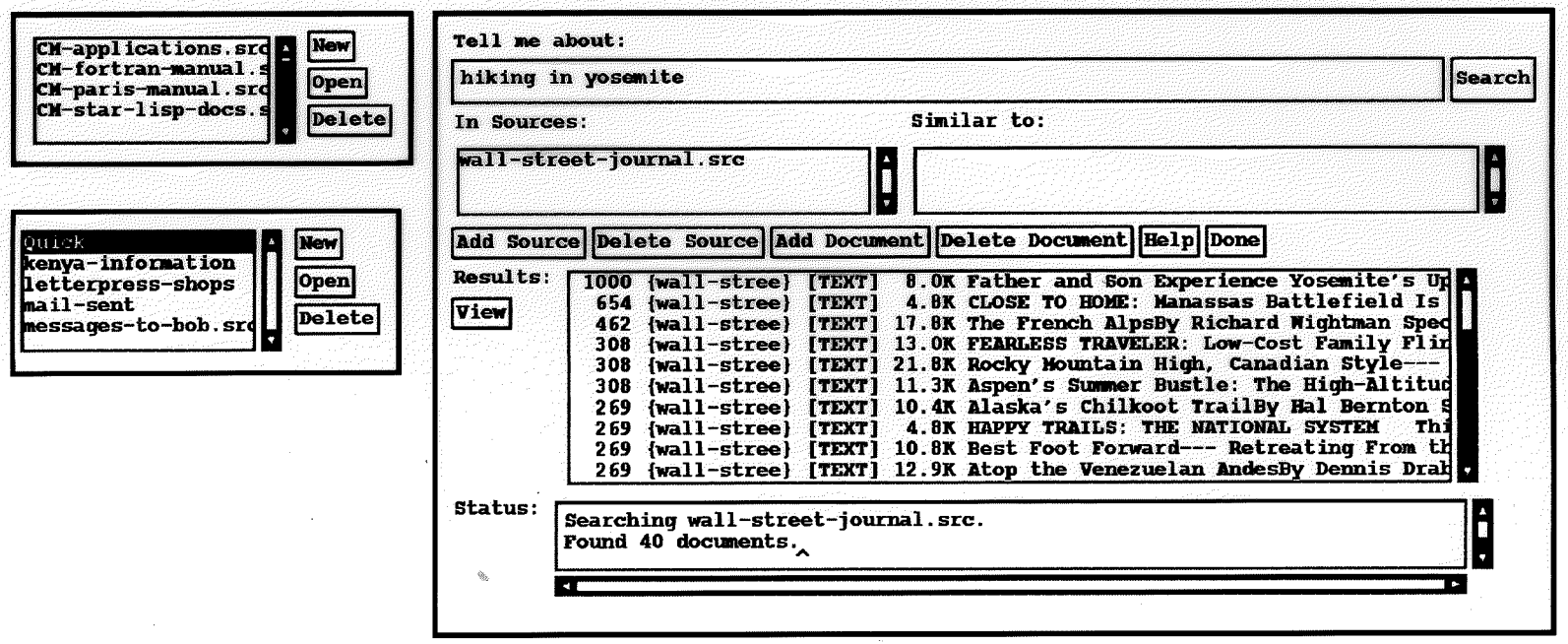
- Metaphor: dynamic folders and conversation
- Windows with drag and drop manipulation
- List of saved servers and questions: Multi-context
- Scores graphically represented
- Sources: alerting, budgets, and confidence
- Document Window: best section, highlighting downloading parts of documents
- Multimedia display

Designer: Harry Morris of Thinking Machines
Use: anonymous FTP from /public/wais/WAIStation*@think.com

WAIStation: Problems

- Directory of servers
- Documents not in original format
- Relevance feedback was not obvious
- 2400 baud was too slow (> 10 sec. search)
- Drag and drop was not obvious to new users
- Alerting was awkward
- Headlines are often not enough context

XWAIS Question Screen



XWAIS Source Screen

Name: world-factbook.src

Server: cmns.think.com

Service: 210

Database: CIA

Cost: q

Units: :free

Maintainer: ephraim@think.com

Description:

Connection Machine WAIS server. Operated between 9AM and 9PM East coast time. The 1990 World Factbook by the CIA whicl good description of every country. The entry for WORLD is al: particularly good.

Save Cancel

XWAIS Document Screen

DJ_doc_id: 0000005648WP BC-07/31/88-D13FATWP
July 31, 1988
Copyright (c) 1988 The Washington Post Co.
California Climbing and hiking
Periodical: WASHINGTON POST: SPORTS, Page d13 (WP)

Headline:
Father and Son Experience Yosemite's Ups and Downs

Climbing Pair Takes California Park's 'Arches' in Stride
By John M. Berry Washington Post Staff Writer
Other than for campers and casual tourists snapping photos of its incredible beauty, Yosemite Valley is a place for the strong, expert climber. The best of them cruise up El Capitan's "nose" in a day or do the 17 pitches of the Royal Arches Route as an easy approach to harder

Add Section | Find Key | Next | Previous | Save To File | Done

Status:

XWAIS: What's New?

- **Interapplication communication for display**
- **Alerting via cron**
- **Quick to implement: 2 months**

Who: Jonathan Goldman of Thinking Machines
Use: anonymous FTP: /public/wais/wais-8*@think.com

XWAIS: Problems

- Multiple processes not centrally controlled
- Hard to install

Rosebud: What's New?

- **Color**
- **Designed for alerting**
- **Newspaper format using first part of documents**
- **Notebook organizer**

Designer: Thomas Erickson and Gitta Salomon
Use: Not public

Rosebud: Problems

- Dealing with multiple servers
- No support for Directory of Servers
- No multimedia
- No relevance feedback of sections

New Problems (revisited)

- Your machine will be spending your money
- The information on your computer will constantly changing
- Browsing online information will become common
- Your machine will know more about you than your spouse

Conclusion

- Beginning of development: infrastructure is stabilizing
- Treat source selection specially
- Alerting interfaces are untested and new

Rosebud Reporter

Rosebud Interactive Response

Rosebud Newspaper

Rosebud Notebook